



# OPEN Contrasting clonal and population genetic structure in two endangered Costa Rican *Vanilla* species of commercial interest

Maria Alejandra Serna-Sánchez<sup>1,2✉</sup>, Adam P. Karremans<sup>1,3</sup>, Diego Bogarín<sup>1,3,4</sup> & Eric J. Fuchs<sup>2,3,5</sup>

Wild relatives of the commercially grown *Vanilla planifolia*, such as *Vanilla odorata* and *V. pompona*, are understudied despite their ecological and agricultural importance. Both species are listed as endangered by the IUCN, but limited research on their genetic diversity hinders effective management. While *Vanilla* species reproduce extensively by cloning, wild populations may retain significant genetic diversity valuable for crop improvement and conservation. To evaluate genetic diversity and structure, we analyzed 146 individuals from 10 *V. pompona* populations and 75 individuals from six *V. odorata* populations in Costa Rica, using 11 and 10 microsatellites, respectively. *Vanilla odorata* showed higher clonality than *V. pompona* (63% vs. 35%). Both species exhibited low to moderate genetic diversity (*V. odorata*:  $H_O = 0.47$ ,  $H_E = 0.29$ ; *V. pompona*:  $H_O = 0.43$ ,  $H_E = 0.41$ ). *Vanilla odorata* populations revealed greater genetic differentiation than *V. pompona* populations (Nei's  $G_{ST} = 0.51$  vs. 0.091), with *V. odorata* also exhibiting isolation by distance ( $R^2 = 0.704$ ,  $p < 0.05$ ). Both species showed fine-scale genetic structure, with higher levels linked to clonality. Gene flow in *V. pompona* benefits from broad distribution and sexual reproduction, while *V. odorata* experiences limited gene flow due to its high clonality and restricted range. Protecting habitats and promoting outcrossing may aid *V. odorata* viability while *V. pompona* serves as an important resource for increasing genetic diversity in cultivated vanilla.

**Keywords** Fine-scale genetic structure, Microsatellites, Pacific region, Population genetics

Approximately 80 million years ago, the Vanilloideae subfamily diverged within Orchidaceae<sup>1</sup>, which includes the genus *Vanilla* Miller. This genus of culinary and cosmetic value includes around 140 species<sup>2</sup>. Among these, only *Vanilla planifolia* Andrews is cultivated on a large scale, with *Vanilla pompona* Schiede and the hybrid *Vanilla × tahitensis* (*V. planifolia* Andrews × *Vanilla odorata* C.Presl) cultivated to a lesser extent<sup>3</sup>. They are cultivated for their fruits rich in the organic compound vanillin<sup>4,5</sup>. *Vanilla planifolia* crops typically have low genetic diversity due to the use of a small number of clones for reproduction<sup>6,7</sup>, which can result in low resilience to environmental changes and biotic factors such as pests and pathogens<sup>8</sup>. Moreover, modern plantations rely on cultivated clones originally introduced from Mesoamerica to the rest of the world during the late 16th century<sup>9</sup>. Certain varieties, such as the high yielding 'Mansa' variety (with more than 75% of fruits retained), are commonly marketed for establishing new plantations, resulting in an industry that relies on a limited genetic base<sup>10</sup>.

Although commercially *V. planifolia* has been the most exploited species, the neotropics have around 41 aromatic *Vanilla* species with the potential for commercial use<sup>11</sup>. In Costa Rica, two wild species with fragrant fruits, *V. odorata* and *V. pompona*, represent critical genetic resources for *Vanilla* cultivation and could potentially contribute to the vanilla extract trade<sup>2</sup>, despite being classified as endangered by the International Union for Conservation of Nature (IUCN). Since 2004, studies on genetic diversity in some *Vanilla* species have been conducted using markers ranging from RAPDs to SNPs. Most of these studies have focused on species-

<sup>1</sup>Centro de Investigación Jardín Botánico Lankester, Universidad de Costa Rica, P.O. Box 302-7050, Cartago, Costa Rica. <sup>2</sup>Programa de Posgrado en Biología, Sistema de Estudios de Posgrado, Universidad de Costa Rica, San Pedro, San José 11501-2060, Costa Rica. <sup>3</sup>Escuela de Biología, Universidad de Costa Rica, Ciudad Universitaria Rodrigo Facio, San Pedro Montes de Oca 11501-2060, Costa Rica. <sup>4</sup>Evolutionary Ecology Group, Naturalis Biodiversity Center, Leiden, The Netherlands. <sup>5</sup>Laboratorio Binacional, UNAM-UCR, Morelia, Michoacán, Mexico. ✉email: maria.serna@ucr.ac.cr

level diversity, with few examining population-level genetic diversity and structure<sup>3,6,8,9,12–17</sup>. In these studies, *V. dilloniana*, *V. imperialis*, and *V. pompona* exhibited higher average heterozygosity compared to *V. planifolia*, *V. odorata*, and *V. × tahitensis*<sup>5,9</sup>. Despite efforts to advance our understanding of *Vanilla* genetics, a significant research gap remains concerning the clonal structure, population genetic diversity, and genetic structure of wild *V. odorata* and *V. pompona* populations. This knowledge gap, essential for effectively managing and conserving these species, will be addressed in this study.

Species with allogamous reproduction and the ability for long-distance seed dispersal are likely to exhibit higher levels of genetic variation and lower levels of population differentiation compared to species that regularly self-fertilize and have limited dispersal of propagules<sup>18</sup>. While certain species within *Vanilla* reproduce naturally via allogamous sexual reproduction or self-pollination, the overall success of natural pollination in *Vanilla* is low, as commonly observed in other food-deceptive orchids. For example, natural fruit set in *V. planifolia* in Puerto Rico and Central America is between 1 and 3%<sup>19</sup>, while Soto Arenas (1999) reported an even lower rate in Mexico: 1 fruit per 100–1000 flowers<sup>22</sup>. This is consistent with the 0.65% fertilization ratio reported by Pemberton et al. (2023) in southern Florida<sup>23</sup> and the 4.9% found by Quezada-Euán et al. (2024) in Yucatán, Mexico<sup>24</sup>. Therefore, vegetative or clonal propagation is the most plausible and economically viable type of reproduction in the genus<sup>25</sup>. This may result in widespread spatial distribution of single genets. In natural environments, *Vanilla* individuals may expand over large areas; for example, an individual of *V. planifolia* was observed to cover areas up to 0.2 hectares<sup>22</sup>.

Although populations of *V. odorata* and *V. pompona* have been documented in Costa Rica, their genetic structure and diversity remain unknown. The primary objective of this project is to study the genetic diversity and clonal structure of *V. pompona* and *V. odorata* populations in Costa Rica using microsatellites. We hypothesize that *V. pompona* exhibits higher genetic diversity than *V. odorata* due to its wider distribution and larger population sizes. However, because *V. odorata* and *V. pompona* can reproduce clonally, we expect low overall genetic diversity, significant fine-scale genetic structure, and significant differences in allele frequencies among populations. Understanding the genetic diversity of these *Vanilla* species will provide baseline information for conservation strategies, such as increasing population connectivity or providing guidelines on best practices for translocating individuals among populations. It will also encourage the use of these species to improve the genetic resources of extensively cultivated *Vanilla*. A comprehensive understanding of the genetic diversity and reproductive biology of *Vanilla* species, including clonality and inbreeding in fragmented and geographically distant populations, is essential to preserve the genetic diversity of *Vanilla* crop wild relatives (CWRs) while simultaneously mitigating the impact of extreme climates on vanilla cultivation.

Results  
Clonal assignment and estimates

Among the 75 *V. odorata* individuals collected in the Osa Peninsula, we identified 26 unique genets across all populations. *Vanilla odorata* has an average clonality of 63%. There is great variation in the level of clonal reproduction among populations, with Miramar and VT having the largest levels of clonality (75%). According to Simpson’s Diversity Index (SDI), the probability of finding two individuals with different genotypes is 25% in Miramar, while in Matapalo it rises to 80%. The average evenness is 0.64, indicating moderate genotypic diversity, as the distribution of genotypes within populations is not completely equitable, but neither is it extremely uneven. Matapalo exhibited the highest evenness (0.93), suggesting a homogeneous frequency distribution of all genotypes. In contrast, VT had the lowest evenness (0.43), indicating that a single genotype is predominant throughout the population (Table 1, Supplementary Fig. S1).

Among the 146 *V. pompona* individuals collected in the Pacific Region of Costa Rica, we identified a total of 93 unique genets. *Vanilla pompona* has significantly less clonality, with an average clonality of 35%. Similar to *V. odorata*, there is significant variation in clonality across populations, with VT exhibiting the highest levels at 75%. According to the Simpson’s Diversity Index, the probability of finding two different genets in VT is 66%, while in Rio Jaris the probability of finding two individuals with different genets is 100%. Average evenness is 0.74, suggesting a balanced and diverse meta-population, with multiple genotypes present in relatively similar proportions (Table 2).

Population	Number of individuals	Genets	Clonality (%)	Simpson’s Diversity Index (SDI)	Evenness
PST	24	8	66.66	0.76 (0.74–0.77)	0.46 (0.43–0.48)
VT	16	4	<b>75.00</b>	0.44 (0.74–0.77)	0.43 (0.38–0.50)
PPT	14	6	57.14	0.75 (0.71–0.78)	0.54 (0.49–0.60)
Miramar	8	2	<b>75.00</b>	0.25 (0.00–0.29)	0.64 (0.66–1.00)
Tigre	8	3	62.50	0.68 (0.57–0.71)	0.82 (0.78–0.98)
Matapalo	5	3	40.00	<b>0.80</b> (0.67–0.83)	<b>0.93</b> (0.89–1.00)
Mean	12.50 ± 2.84	4.33 ± 0.91	62.71 ± 5.36	0.61 ± 0.09	0.64 ± 0.08

**Table 1.** Clonal diversity estimates for *V. odorata* populations in Southern Costa Rica. Number of sampled individuals, number of unique genets, clonal percentage, Simpson’s diversity index, and evenness for each population, along with mean ± standard error values across all populations. Jackknife estimates for SDI and evenness are provided in parentheses. Values in bold indicate the highest within each metric, for comparative and visual interpretation purposes only.

Population	Number of individuals	Genets	Clonality (%)	Simpson's Diversity Index (SDI)	Evenness
STH	20	17	15.00	<b>0.98</b> (0.98–0.99)	<b>0.91</b> (0.90–0.92)
VT	20	5	<b>75.00</b>	0.66 (0.63–0.67)	0.54 (0.54–0.61)
PST	12	6	50.00	0.82 (0.78–0.86)	0.67 (0.61–0.75)
Los Chocuacos	16	9	43.75	0.87 (0.85–0.89)	0.59 (0.59–0.64)
Clavera	18	14	22.22	<b>0.97</b> (0.96–0.98)	0.83 (0.82–0.90)
Pavona	18	13	27.77	0.95 (0.94–0.96)	0.73 (0.72–0.82)
Garabito	8	4	50.00	0.82 (0.76–0.86)	0.89 (0.82–0.96)
Rio Jaris	8	8	0.00	<b>1.00</b> (1.00–1.00)	<b>1.00</b> (1.00–1.00)
San Rafael	18	12	33.33	0.92 (0.91–0.93)	0.64 (0.64–0.69)
El Hacha	8	5	37.50	0.79 (0.71–0.86)	0.64 (0.65–0.75)
Mean	14.60 ± 1.60	9.30 ± 1.41	35.45 ± 6.64	0.88 ± 0.03	0.74 ± 0.05

**Table 2.** Clonal diversity estimates for *Vanilla pompona* populations across the Pacific region of Costa Rica. Number of sampled individuals, number of unique genets, clonal percentage, Simpson's diversity index (SDI), and evenness for each population, along with mean ± standard error values across all populations. Jackknife estimates for SDI and evenness are provided in parentheses. Values in bold indicate the highest within each metric, for comparative and visual interpretation purposes only.

Furthermore, our results show that the average maximum clonal patch size in *V. odorata* was 36.2 (± 13.9) meters (Supplementary Fig. S1). For *V. pompona*, the average maximum spread of clones was 28.4 (± 12.5) meters (Supplementary Fig. S2). An intermixed pattern of genets was observed in both species, with different clones and genets coexisting in the same area. However, in *V. odorata* populations such as Miramar, VT, and PST, a few genets dominate, creating large clonal patches that reach up to 10 m in Miramar and at least 74 m in VT (Supplementary Fig. S1). A similar pattern was observed in *V. pompona*, although with greater clonal diversity (SDI) within populations. Only a few populations had large clones, such as VT, which had two genets that extended up to 32 m, and Los Chocuacos, where one genet extended up to 31 m. In the case of El Hacha, the same genet was identified in two individuals separated by up to 112 m.

### Genetic diversity

All markers in both species deviated from Hardy-Weinberg equilibrium (HWE), as expected for clonal species. Across the 10 loci analyzed in *V. odorata*, the number of alleles ranged from two (mVplCIR047) to six (mVhuCIR07). In *V. pompona*, the 11 loci analyzed exhibited a broader range, with three to 15 alleles, with mVhuCIR07 being the most polymorphic. Our estimates of genetic diversity were similar between populations within each species, with few differences between analyses including and excluding clones, except possibly in  $F_{IS}$ . Genetic diversity estimates were slightly influenced by sample size. For example, in *V. odorata* in Miramar, we found the lowest diversity metrics because it is the population with the smallest sample size. The two populations with the fewest genets (VT and Garabito) also had the lowest genetic diversity estimates in *V. pompona*. *Vanilla odorata* shows lower diversity metrics than *V. pompona*, with fewer alleles ( $A = 1.70$  vs. 2.76), lower allelic richness ( $Ar = 1.57$  vs. 2.30), and a smaller effective number of alleles ( $Ae = 1.57$  vs. 2.09). Although *V. pompona* has higher expected heterozygosity ( $He = 0.40$ , 95% CI = 0.20–0.58) compared to *V. odorata* ( $He = 0.27$ , 95% CI = 0.07–0.45), confidence intervals overlap. There is no clear geographic pattern in the distribution of genetic diversity. Inbreeding in *V. pompona* was low and not significant ( $F_{IS} = -0.03$ ,  $SE = 0.19$ ), while a significant heterozygote excess was found in *V. odorata* ( $F_{IS} = -0.67$ ,  $SE = 0.17$ ) (Tables 3 and 4).

### Genetic structure and isolation by distance

We found significant genetic structure at the population level in both species:  $G_{ST} = 0.514$  (95% CI = 0.500–0.529) for *V. odorata*; and  $G_{ST} = 0.091$  (95% CI = 0.075–0.108) for *V. pompona*. Structure was significantly higher in *V. odorata*. The DAPC analysis of *V. odorata* shows a clear separation of populations in the inner part of the peninsula and those populations in Piro, forming a separate cluster (Fig. 1A). We found significant IBD for *V. odorata* ( $r_{IC} = 0.704$ ,  $p < 0.05$ ;  $r_{EC} = 0.737$ ,  $p < 0.05$ ; Supplementary Fig. S3).

In *V. pompona*, the DAPC analysis revealed a very low level of genetic structure, resulting in most populations clustering together. Only two populations, STH and VT, stand out as genetically distinct. Populations from the South Pacific–Térraba (Clavera and Los Chocuacos), Central Pacific–Carara (Pavona and Garabito), North Pacific (El Hacha), and PST all form a central cluster, reflecting similar allele frequencies (Fig. 1B, Supplementary Fig. S6). Additionally, San Rafael and Rio Jaris form a separate but closely related cluster to the main group (Fig. 1B, Supplementary Fig. S6). No significant IBD was found among *V. pompona* populations ( $r_{IC} = 0.138$ ,  $p = 0.208$ ;  $r_{EC} = 0.056$ ,  $p = 0.275$ ; Supplementary Fig. S4) or at the regional level ( $r_{IC} = 0.351$ ,  $p = 0.187$ ;  $r_{EC} = 0.533$ ,  $p = 0.123$ ; Supplementary Fig. S5).

### Fine scale spatial genetic structure

For both species, pairwise relatedness decreased with increasing pairwise distances between individuals (Figs. 2 and 3). The point at which the correlogram crossed the X-axis allowed us to calculate the average patch size<sup>26,27</sup>. In *V. odorata*, the average patch size was estimated at around 4120 m for both datasets (Fig. 2). For *V.*

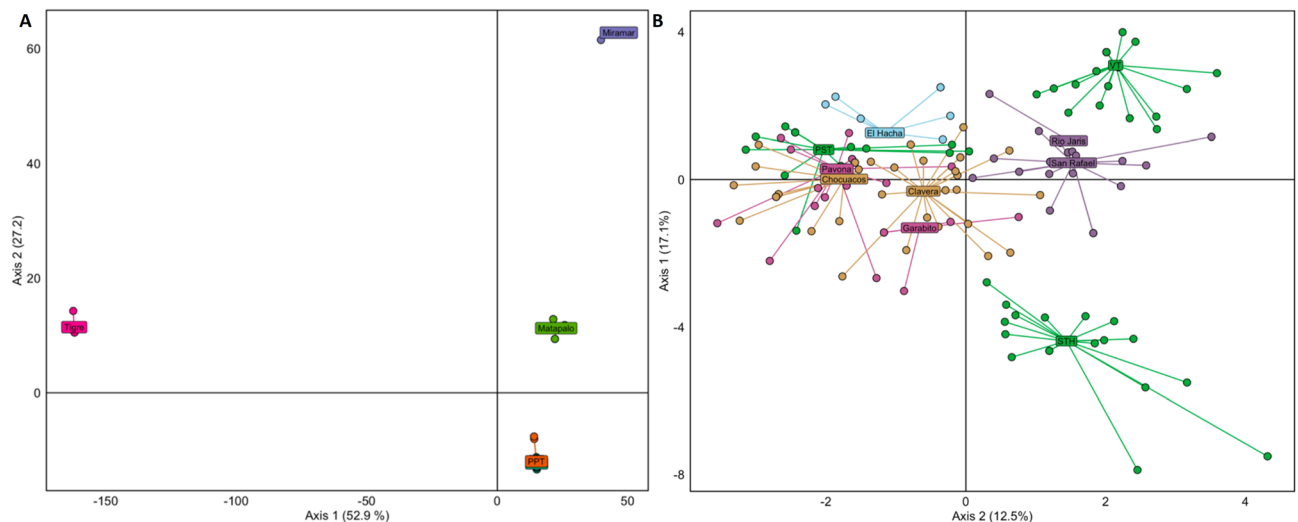
Population	Dataset	N	A	Ar	Ae	Ho	He	F <sub>IS</sub>
PST	IC	24	1.8 (0.25)	1.72 (0.21)	1.70 (0.21)	0.55 (0.15)	0.32 (0.09)	-0.76* (0.14)
	EC	8	1.8 (0.25)	1.62 (0.18)	1.67 (0.20)	0.52 (0.15)	0.32 (0.09)	-0.68* (0.14)
VT	IC	16	1.9 (0.23)	1.76 (0.21)	1.71 (0.21)	0.58 (0.16)	0.33 (0.09)	-0.71* (0.14)
	EC	4	1.9 (0.23)	1.68 (0.16)	1.64 (0.17)	0.55 (0.15)	0.34 (0.08)	-0.59* (0.19)
PPT	IC	14	1.9 (0.23)	1.79 (0.22)	1.73 (0.22)	0.57 (0.15)	0.33 (0.09)	-0.67* (0.14)
	EC	6	1.9 (0.23)	1.68 (0.18)	1.73 (0.21)	0.57 (0.15)	0.35 (0.09)	-0.62* (0.15)
Miramar	IC	8	1.4 (0.16)	1.36 (0.15)	1.31 (0.15)	0.31 (0.15)	0.16 (0.07)	-0.75* (0.16)
	EC	2	1.4 (0.16)	1.40 (0.16)	1.36 (0.15)	0.35 (0.15)	0.20 (0.08)	-0.75* (0.16)
Tigre	IC	8	1.5 (0.22)	1.48 (0.21)	1.39 (0.16)	0.34 (0.15)	0.20 (0.08)	-0.68 (0.20)
	EC	3	1.5 (0.22)	1.46 (0.20)	1.46 (0.19)	0.33 (0.15)	0.23 (0.10)	-0.60 (0.25)
Matapalo	IC	5	1.6 (0.22)	1.60 (0.22)	1.56 (0.20)	0.48 (0.16)	0.26 (0.09)	-0.83* (0.07)
	EC	3	1.6 (0.22)	1.55 (0.19)	1.54 (0.19)	0.47 (0.16)	0.27 (0.09)	-0.77* (0.10)
Mean	IC	75	1.7 (0.22)	1.62 (0.20)	1.57 (0.19)	0.47 (0.15)	0.27 (0.09)	-0.73* (0.14)
	EC	26	1.7 (0.22)	1.57 (0.18)	1.57 (0.19)	0.47 (0.15)	0.29 (0.09)	-0.67* (0.17)

**Table 3.** Genetic diversity estimates for six populations of *V. odorata* in Costa Rica, based on a dataset including clones (IC) and excluding clones (EC). *N* sample size, *A* average number of alleles per locus, *Ar* allelic richness, *Ae* effective number of alleles, *Ho* observed heterozygosity, *He* expected heterozygosity, *F<sub>IS</sub>* inbreeding coefficient. Mean values  $\pm$  (SE). \*Indicates significant *F<sub>IS</sub>* ( $P < 0.01$ ) based on 1000 bootstrap replicates.

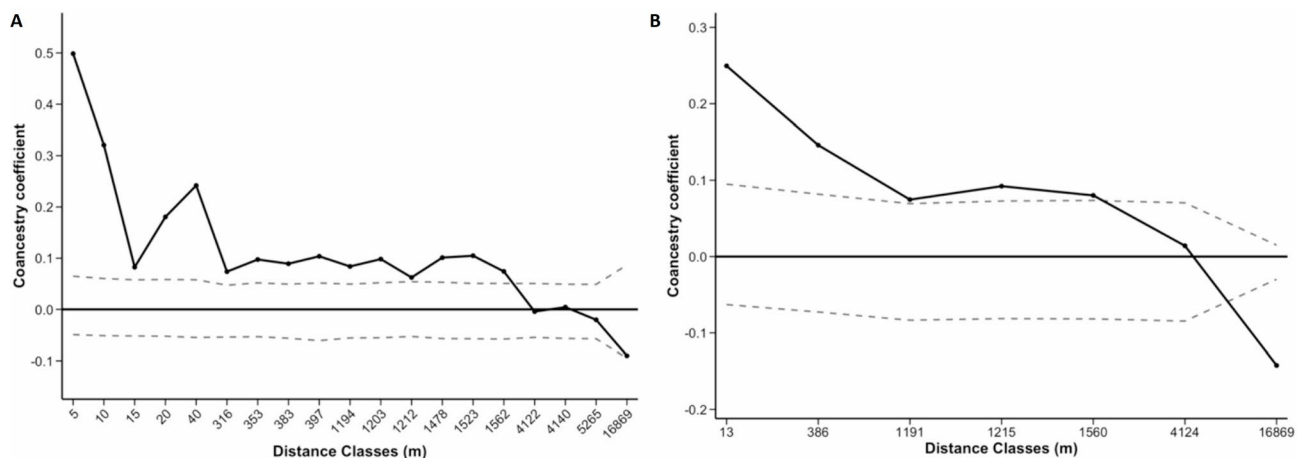
Population	Dataset	N	A	Ar	Ae	Ho	He	F <sub>IS</sub>
STH	IC	20	2.91 (0.64)	2.48 (0.43)	2.03 (0.35)	0.44 (0.12)	0.40 (0.08)	-0.05 (0.16)
	EC	17	2.91 (0.64)	2.26 (0.34)	2.10 (0.38)	0.44 (0.12)	0.42 (0.08)	-0.04 (0.16)
VT	IC	20	2.09 (0.28)	1.99 (0.27)	1.73 (0.28)	0.40 (0.12)	0.31 (0.08)	-0.24 (0.14)
	EC	5	1.91 (0.28)	1.86 (0.27)	1.59 (0.22)	0.33 (0.12)	0.30 (0.08)	-0.06 (0.19)
PST	IC	12	2.55 (0.34)	2.24 (0.26)	1.75 (0.19)	0.55 (0.13)	0.36 (0.07)	-0.37* (0.16)
	EC	6	2.45 (0.37)	2.20 (0.29)	1.85 (0.23)	0.52 (0.13)	0.39 (0.08)	-0.28* (0.15)
Los Chocucos	IC	16	2.82 (0.6)	2.46 (0.41)	2.16 (0.35)	0.53 (0.14)	0.46 (0.07)	-0.12 (0.23)
	EC	9	2.82 (0.60)	2.35 (0.36)	2.21 (0.42)	0.50 (0.12)	0.46 (0.07)	-0.06 (0.22)
Clavera	IC	18	3.45 (0.78)	3.00 (0.60)	2.54 (0.67)	0.47 (0.12)	0.46 (0.08)	0.14 (0.20)
	EC	14	3.45 (0.78)	2.58 (0.41)	2.48 (0.61)	0.48 (0.12)	0.47 (0.07)	0.13 (0.21)
Pavona	IC	18	3.82 (0.84)	3.06 (0.55)	2.52 (0.53)	0.47 (0.11)	0.48 (0.08)	0.05 (0.13)
	EC	13	3.73 (0.84)	2.69 (0.40)	2.55 (0.52)	0.43 (0.11)	0.50 (0.08)	0.16 (0.15)
Garabito	IC	8	2.18 (0.44)	2.17 (0.44)	1.97 (0.40)	0.35 (0.14)	0.36 (0.10)	-0.04 (0.25)
	EC	4	2.18 (0.44)	2.18 (0.44)	1.94 (0.38)	0.36 (0.14)	0.39 (0.10)	-0.01 (0.25)
Rio Jaris	NC	8	2.73 (0.69)	2.26 (0.45)	2.12 (0.48)	0.43 (0.13)	0.37 (0.10)	-0.11 (0.19)
San Rafael	IC	18	2.82 (0.55)	2.42 (0.42)	1.90 (0.33)	0.35 (0.12)	0.35 (0.09)	0.14 (0.18)
	EC	12	2.82 (0.55)	2.21 (0.33)	1.89 (0.29)	0.37 (0.12)	0.38 (0.08)	0.10 (0.18)
El Hacha	IC	8	2.55 (0.43)	2.50 (0.42)	1.99 (0.27)	0.45 (0.13)	0.40 (0.10)	-0.14 (0.17)
	EC	5	2.55 (0.43)	2.42 (0.39)	2.12 (0.33)	0.45 (0.13)	0.43 (0.11)	-0.11 (0.16)
Mean	IC	146	2.79 (0.56)	2.50 (0.44)	2.07 (0.39)	0.44 (0.13)	0.40 (0.09)	-0.07 (0.18)
	EC	93	2.76 (0.56)	2.30 (0.37)	2.09 (0.39)	0.43 (0.12)	0.41 (0.09)	-0.03 (0.19)

**Table 4.** Genetic diversity estimates for 10 populations of *V. pompona* in Costa Rica, based on a dataset including clones (IC) and excluding clones (EC). *N* sample size, *A* average number of alleles per locus, *Ar* allelic richness, *Ae* effective number of alleles, *Ho* observed heterozygosity, *He* expected heterozygosity, *F<sub>IS</sub>* inbreeding coefficient. Mean values  $\pm$  (SE). \*Indicates significant *F<sub>IS</sub>* ( $P < 0.01$ ) based on 1000 bootstrap replicates.

*pompona*, average patch size was approximately 92 m when clones were included and 283 m when clones were excluded (Fig. 3). Autocorrelation between *V. odorata* individuals separated by less than 5 m, including ramets ( $r = 0.498$ ; Fig. 2A), was significantly different from zero and twice as large as for our data set with only genets ( $r = 0.249$ ; Fig. 2B). In *V. pompona*, autocorrelation between individuals separated by less than 3 m, including ramets ( $r = 0.250$ ; Fig. 3A), was significant and slightly higher than for individuals separated by less than 4 m in the genets subset ( $r = 0.190$ ; Fig. 3B). However, spatial autocorrelation remains significant and high in *V.*



**Fig. 1.** Discriminant analysis of principal components (DAPC) (A) clustering of six populations of *V. odorata*, including clones (IC). (B) clustering of ten populations in *V. pompona*, including clones. Colors denote the geographical allocation of *V. pompona* populations (clusters) to designated regions: Green = South Pacific - Peninsula; Yellow = South Pacific - Térraba; Pink = Central Pacific - Carara; Blue = North Pacific; Purple = Central Pacific - Puriscal.



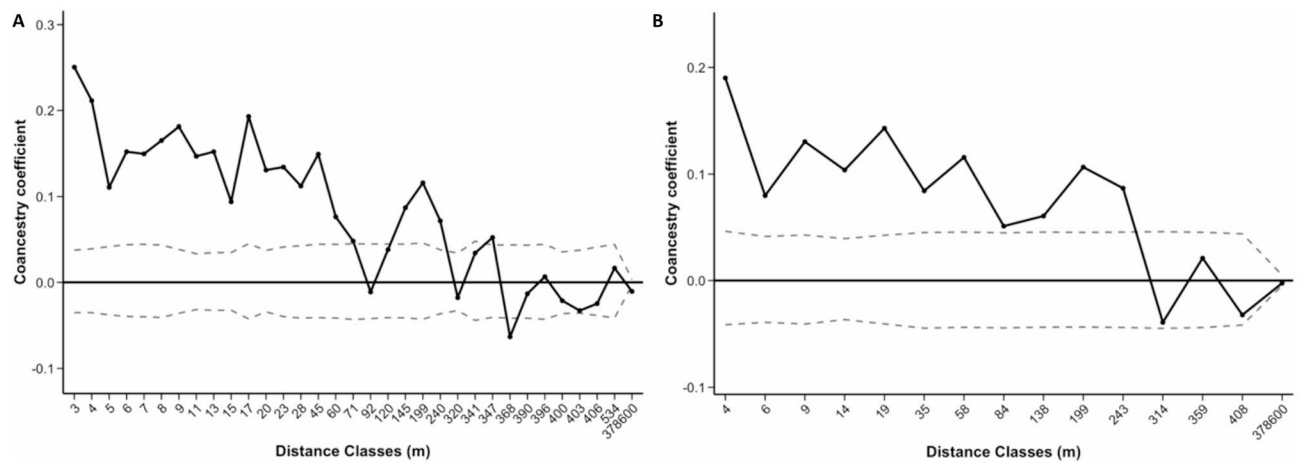
**Fig. 2.** Correlograms showing the spatial autocorrelation of the coancestry Loiselle coefficient (Loiselle et al., 1995) as a function of distance for *V. odorata* populations: (A) dataset including clones, (B) dataset excluding clones. Dashed lines represent the lower and upper 95% confidence limits for the null hypothesis of no spatial genetic structure.

*odorata* up to 1560 m, and up to 60 m and 243 m in *V. pompona* when clones are included (IC) and excluded (EC), respectively (Figs. 2 and 3). Loiselle's coancestry coefficient was consistently higher for the ramet dataset than for the genet dataset, indicating a strong influence of clones in shaping spatial genetic structure in the closest distance classes. When FSGS analyses were performed by population for each species, similar patterns were observed (Supplementary Figs. S7 and S8). However, individual autocorrelation was lower because larger distance classes were needed due to the limited number of individuals within populations.

## Discussion

In our studies, the clonal nature of *V. odorata* and *V. pompona* has proven to be a critical factor in understanding the amount and distribution of genetic diversity. In comparing *V. odorata* and *V. pompona*, our study highlights sharp differences in clonality, geographic distribution, genetic diversity, and structure. In Costa Rica, while *V. odorata* is more clonal and has a restricted distribution, primarily confined to the Osa Peninsula, *V. pompona* has a broader distribution along Costa Rica's Pacific coast. As expected, *V. pompona* showed higher levels of genetic diversity compared to *V. odorata*, likely due to its wider distribution, reduced clonality, and increased opportunities for sexual reproduction and gene flow. This is confirmed by *V. pompona*'s lack of inbreeding and





**Fig. 3.** Correlograms showing the spatial autocorrelation of the coancestry Loiselle coefficient (Loiselle et al., 1995) as a function of distance for *V. pompona* populations: **(A)** dataset including clones, **(B)** dataset excluding clones. Dashed lines represent the lower and upper 95% confidence limits for the null hypothesis of no spatial genetic structure.

low levels of spatial genetic structure in contrast with *V. odorata*. The restricted and clonal nature of *V. odorata* populations is consistent with our predictions, as it results in heterozygote excess due to its limited genetic diversity and predominantly vegetative propagation.

In our results, repeated genotypes were considered products of vegetative reproduction, particularly guerrilla-type growth, where ramets are widely spaced and rapidly spread into new areas<sup>28</sup>. Vegetative or clonal reproduction in orchids often correlates with reduced genetic diversity, as observed in *Vanilla planifolia*<sup>6,7</sup>. Similarly, the orchid *Pelatantheria scolopendrifolia* (Makino) Aver. had moderate to low levels of genetic diversity, attributed primarily to small population sizes, demographic events, clonal reproduction, and limited gene flow<sup>29</sup>. Clonal architecture, which influences the spatial distribution of ramets, can significantly impact geitonogamy<sup>30</sup>. In populations of *V. odorata*, such as Miramar, Tigre, VT, and PST, a clumped distribution of ramets was observed. In *V. odorata*, this clustering will likely lead to a higher degree of geitonogamy<sup>30</sup> and may contribute to the significant genetic structure observed among populations. Conversely, *V. pompona* individuals are typically more intermingled (except in VT), which may lead to higher outcrossing rates<sup>30</sup>. A mixed distribution of genets can improve outcrossing by increasing the availability of diverse pollen from multiple genets<sup>30</sup>. This pattern is commonly observed in populations of *V. pompona* (Supplementary Fig. S2).

The near-zero inbreeding coefficient observed in *V. pompona* ( $F_{IS} = -0.03$ ) is in direct contrast with high inbreeding values reported for *V. mexicana* ( $F_{IS} = 0.74$ ), where autogamy is prevalent<sup>16</sup>. Thus, the higher levels of genetic diversity observed in *V. pompona* are most likely due to relatively higher levels of sexual reproduction compared to clonal growth. In contrast, *V. odorata*'s heterozygote excess ( $F_{IS} = -0.67$ ) may be attributed to its high clonal reproduction rates that maintain large fractions of heterozygous ramets within populations. This pattern is also consistent with bottleneck or founder effects, where guerrilla propagation of clones may lead to the widespread establishment of heterozygous genotypes<sup>31</sup>. Clonal reproduction may maintain heterozygosity or even increase it by mutation over generations<sup>32,33</sup>.

The genetic diversity observed in *Vanilla odorata* and *V. pompona* aligns with patterns observed in other orchids, with diversity levels influenced by the degree of clonality. In *V. odorata*, smaller and highly clonal populations, such as Miramar, Tigre, and Matapalo, showed notably lower heterozygosity values compared to larger, less clonal populations like PPT and PST. Similarly, in *V. pompona*, the more clonal populations VT, PST, and Garabito had reduced genetic diversity, whereas larger populations like STH, Clavera, and La Pavona exhibited higher genetic diversity. Previous studies have shown that in clonal orchids, such as *Cypripedium calceolus* L., populations that rely more heavily on clonal reproduction tend to have lower genetic diversity due to limited genetic recombination compared to populations that favor sexual reproduction<sup>34</sup>. When compared with other *Vanilla* species, *V. odorata* exhibits low to moderate genetic diversity, which aligns with findings for *V. planifolia*, another highly clonal species with a deceptive pollination system<sup>21</sup>. *Vanilla pompona* has higher levels of genetic diversity, similar to African *Vanilla* species (*V. madagascariensis*, *V. decaryana*, *V. humblotii*, *V. perrieri*, and *V. bosseri*), which combine both clonal and sexual reproductive strategies but tend to have lower clonal reproduction<sup>15,35</sup>. Like *V. pompona* populations in Costa Rica, *V. madagascariensis*, *V. decaryana*, and *V. perrieri* also tend to have larger populations and broader distributions across Madagascar, contributing to their higher genetic diversity. Together, these results underscore how reproductive strategy and population size can influence genetic diversity, with clonality often associated with reduced diversity due to limited genetic recombination and smaller effective population sizes<sup>36,37</sup>.

Gene flow within Orchidaceae has been shown to occur over distances of up to 250 km<sup>38</sup>, but greater population differentiation can arise when intermediate populations are scarce, limiting gene flow between distant populations<sup>38</sup>. The genetic structure in *Vanilla odorata* and *V. pompona* likely reflects the combined influence of limited gene flow through pollen and seed dispersal, shaped by distinct pollination and dispersal mechanisms

in each species. In *V. odorata*, a higher population structure is observed, likely due to pollination through food deception by both male and female *Euglossa* bees<sup>21</sup>. In deceptive systems, pollinators receive no reward (such as food or fragrance), so they quickly learn to avoid these flowers, resulting in less frequent visits and, consequently, fewer opportunities for effective pollination<sup>39</sup>. Limited pollen dispersal in *V. odorata*, combined with short-range seed dispersal by fragrance and pulp collection by *Euglossa* bees and *Trigona fulviventrtris* (Meliponini)<sup>40</sup>, likely contributes to the Isolation by Distance (IBD) pattern and strong genetic clustering, particularly in geographically isolated populations like Tigre, Miramar, and Matapalo.

Conversely, *V. pompona*, which is distributed along the Pacific coast, exhibits lower genetic structure despite its broad range. This pattern can be explained by the stepping-stone theory, where populations connected by intermediate groups maintain genetic connectivity across larger distances, even if direct gene flow between distant populations is limited<sup>41,42</sup>. In *V. pompona*, *Eulaema cingulata* was identified as the main pollinator, engaging in floral scent collection and nectar searching<sup>43</sup>. This behavior facilitates pollen removal and attracts bees from long distances, potentially increasing gene flow<sup>43</sup>. In addition, *V. pompona* benefits from effective seed dispersers like the spiny rat (*Proechimys semispinosus*) and agouti (*Dasyprocta punctata*), which have extensive home ranges (up to 1,258 m during the rainy season) and play a significant role in maintaining genetic connectivity across Central American forests<sup>44</sup>. These dispersers help spread seeds over greater distances, supporting high genetic connectivity in *V. pompona* despite geographical separation. Overall, these findings highlight how biological interactions, as well as clonality, population size and gene flow, influence genetic structure in *V. odorata* and *V. pompona*.

Interspecific gene flow via hybridization may also contribute to genetic diversity and structure estimates in our samples, as *V. pompona* and *V. odorata* are sympatric. Although natural hybrids between *V. pompona* and *V. odorata* may exist, their occurrence under natural conditions is highly unlikely due to the specificity in plant-pollinator relationships and the specialized pollination strategies in *Vanilla*. As previously stated, the pollinators of *V. pompona* are male *Eulaema*, a large bee that is attracted to the flowers by floral fragrances, which are the main reward<sup>43</sup>. Conversely, the deceptive *V. odorata* is likely pollinated by smaller *Euglossa* or Meliponini bees, analogous to the pollination of the florally similar *V. planifolia*<sup>21</sup>. Furthermore, in natural conditions, interspecific pollen transfer is highly unlikely due to the lack of flowering overlap between both species<sup>43,45</sup>. Considering the morphological differences of pollinators and the lack of synchrony in flowering phenology, natural hybridization between *V. odorata* and *V. pompona* is unlikely.

Fine-scale genetic structure (FSGS) in plant populations is defined as the non-random distribution of genotypes at a local scale, which can be caused by limited seed dispersal, clonality, or environmental factors<sup>46</sup>. In *V. odorata*, the strong genetic structuring observed when clones were included, suggests that clonality plays a dominant role in shaping local genetic patterns, as reported in previous studies<sup>14,36</sup>. However, after removing clones, the genetic structure revealed a pattern consistent with full-sibling relationships<sup>47</sup>, indicating that seed dispersal in sexually reproduced progeny is highly localized. In *V. pompona*, a similar but less clear pattern was observed. The moderate relatedness observed when clones were included, indicates a mixture of clonal individuals, sexually produced progenies, and unrelated individuals. Once clones were excluded, relatedness decreased significantly but remained significant up to 240 m, indicating large clusters of related individuals composed of full- and half-siblings within *V. pompona* populations. A similar pattern was observed in *V. humblotii* Rchb.f., where FSGS was greater for the ramets (datasets with clones) than for the genets alone, suggesting that, like in *V. odorata* and *V. pompona*, the repeated multilocus genotypes (clones) play a significant role in defining a strong spatial genetic structure, particularly in the first distance classes<sup>14</sup>.

Despite *V. pompona*'s potential for long-distance gene dispersal, many seeds may remain in close proximity, leading to local FSGS (see Fig. 3). Studies regarding the seed dispersal of Neotropical *Vanilla* suggest, that although potentially being able to have long-distance dispersal, most seeds are likely being dispersed relatively short distances from the plant<sup>40,48</sup>. Specifically, *Vanilla pompona* fruits are indehiscent, naturally dropping below the plant when mature. The fruit is then consumed fully or in part by mammals, especially rodents, which can disperse the seeds up to 18 h after consumption, usually within its home range<sup>40</sup>. In the case of *Vanilla odorata* fruits are dehiscent, naturally exposing the seeds that are displaced by fragrance collecting male orchid bees, causing seeds to drop directly below the fruit<sup>40</sup>. This localized seed dispersal, coupled with the behavior of seed vectors, reinforces the development of fine-scale genetic structure in both *V. pompona* and *V. odorata*.

Our findings highlight that both population size and clonality significantly shape genetic diversity in *V. odorata* and *V. pompona*. Additionally, clonality increases genetic structuring, particularly in small, isolated populations. From a conservation perspective, our study underscores the importance of maintaining larger population sizes and reducing isolation to support genetic diversity, especially in clonal species like *V. odorata*, where genetic diversity is already limited. Preserving diverse and interconnected habitats for *V. pompona* could enhance gene flow, helping to maintain population sizes and resilience. For *V. odorata*, focusing on reducing genetic bottlenecks by protecting diverse habitats and promoting occasional outcrossing could mitigate genetic drift and support long-term viability. These strategies are essential for conserving these unique genetic resources and the adaptive potential of both species in Costa Rica's tropical forests. To achieve this goal, integrated conservation strategies recommended by Watteyn et al. (2020) include protecting forest areas where wild vanilla species naturally occur, promoting reforestation of biological corridors, cultivating vanilla alongside native tree species with economic benefits, and implementing payments for ecosystem services to incentivize farmers and landowners to conserve native vanilla populations while adopting sustainable agricultural practices<sup>11</sup>.

Moreover, our results suggest that *V. pompona* may be an important resource to enhance the genetic diversity of cultivated vanilla through breeding strategies that include hybridization with *V. planifolia*. This approach could promote sexual reproduction while maintaining high genetic diversity. The broad distribution and reduced genetic structure of *V. pompona* make it a valuable resource from which individuals can be selected from various sites. Preliminary studies indicate that wild *V. pompona* populations possess desirable traits for

commercial plantations, such as increased resistance to pathogens like *Fusarium oxysporum* f. sp. *vanillae* and increased drought tolerance, compared to cultivated *V. planifolia*<sup>49,50</sup>. Menchaca et al. (2011) confirmed the feasibility of generating hybrids between *V. planifolia* and *V. pompona*, which exhibited high germination rates and remarkable hybrid vigor<sup>51</sup>. On the other hand, it is crucial to note that when collecting individuals of *V. odorata* from patches of 30–40 m, there is a high likelihood that they belong to the same clone. This highlights the importance of considering spatial genetic structure when collecting and using wild populations for breeding programs. These findings emphasize the importance of conserving multiple populations to preserve the overall genetic diversity of the species.

While this study employed microsatellites and included a robust sampling effort, certain limitations should be acknowledged. The limited number of populations and markers, particularly for *Vanilla odorata*, a rare species confined to just a few localities in Costa Rica, may influence the scope of our findings. For future research, we recommend expanding marker types and genomic coverage, such as using SNPs, as well as incorporating nuclear and organellar markers like plastomes to provide complementary insights into the species' evolutionary history. Additionally, including more populations across a broader geographic range and integrating landscape genetics would help to further explore how geographic and environmental factors shape genetic variation.

## Methods

### Study species

*Vanilla* is a pantropical genus consisting of around 140 species of hemiepiphytic and epiphytic vines within the Orchidaceae family<sup>2</sup>. *Vanilla odorata* and *V. pompona* are distributed from Mexico to Peru and Brazil<sup>45</sup>, with both species present in Costa Rica's Osa Conservation Area (ACOSA)<sup>11</sup>. *Vanilla pompona* also occurs along the Pacific slope of Costa Rica. Their flowers consist of three sepals, two lateral petals, and a large trumpet-shaped petal known as the labellum, which is partially fused to the column. This leads to a single ventral stigma, separated by the rostellum from the anther, which holds two pollen masses<sup>45</sup>. *Vanilla pompona* blooms in Costa Rica from January to February<sup>43</sup>, with flowers measuring 6 to 8 cm in length and 7.5 to 10 cm in diameter. The successive inflorescences have one to two open flowers at a time, which are ephemeral, lasting only one day and wilting after noon; they are pale yellow with a striking yellow-orange labellum and a strong mint-like fragrance<sup>45</sup>. In contrast, *Vanilla odorata* flowers bloom in April and May<sup>45</sup>, and are 8 cm long and 7 cm wide<sup>45</sup>. They feature successive inflorescences with one to two open flowers at a time, ephemeral from 7:00 a.m. to 4:00 p.m., with pale greenish-white tepals and a greenish-white labellum, and have a mild, fresh fragrance<sup>45</sup>.

### Sample collection

The localities cited by Karremans et al. (2020) for *V. odorata* and *V. pompona* were visited and sampled in the current study, with the materials collected being highly representative of the known populations of both species in the country. *Vanilla* plants used for this study were identified by A.P. Karremans, following the circumscription of Karremans et al. (2020). Representative vouchers for each study site for both species are kept as living plants or liquid specimens at Centro de Investigación Jardín Botánico Lankester (JBL), in Cartago, Costa Rica. A complete list of voucher numbers is provided in Supplementary Table S1. We selected 10 *V. pompona* populations along the Pacific slope of Costa Rica (Fig. 4). For *V. odorata*, we collected samples from six populations located in the Osa Peninsula (Fig. 4). We sampled between 8 and 20 individuals from each of the 10 *V. pompona* populations, totaling 146 individuals; these individuals were grouped into five regions (See Fig. 1). For *V. odorata*, we sampled from 5 to 24 individuals from each of the six populations. Each individual was georeferenced, and we collected one to two leaves per individual, which were stored in silica gel and processed within three days at the Laboratorio de Ecología Molecular of the School of Biology at Universidad de Costa Rica. All procedures complied with Costa Rican laws, and research permits were approved by the Research Committee of the School of Biology at the University of Costa Rica under project approval code C0049 from the Vice-Rectorate of Research.

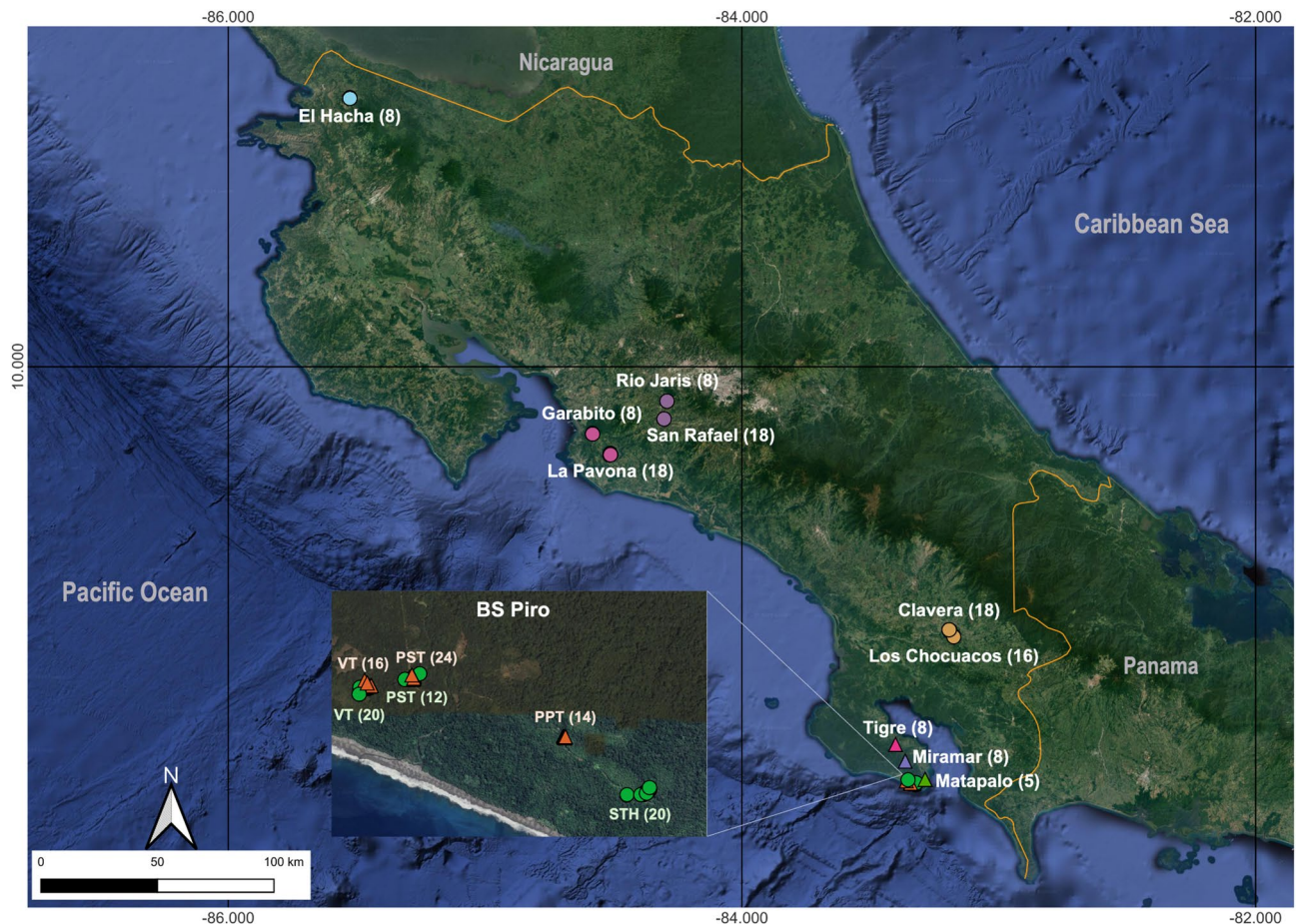
### DNA extraction and SSR amplification

We extracted total DNA from leaf tissue using the modified CTAB protocol from Doyle and Doyle (1990)<sup>52</sup>. Dilutions were prepared to achieve a final concentration of 20 ng/μL for PCR reactions. Microsatellites (SSRs) previously proposed in the literature for other *Vanilla* species were used in this study<sup>6,17</sup>. A total of 28 SSR markers were tested for transferability, of which 16 successfully amplified in *V. pompona* and 15 in *V. odorata*. Among these, 12 markers were polymorphic for each species. However, one marker was excluded from the analysis in *V. pompona* and two in *V. odorata* due to missing data. Therefore, eleven microsatellite loci were used for *V. pompona*: mVhuCIR07, mVhuCIR08, mVhuCIR10, mVhuCIR11, mVroCIR01, mVroCIR05; mVplCIR015, mVplCIR016, mVplCIR019, mVplCIR026, mVplCIR031; while 10 microsatellite loci were used for *V. odorata*: mVhuCIR07, mVhuCIR08, mVroCIR01, mVroCIR03, mVroCIR05; mVplCIR05, mVplCIR025, mVplCIR026, mVplCIR031, mVplCIR047. The forward primer of each primer pair was fluorescently labeled at the 5' end. PCR thermal profiles and primer mixes are described in Supplementary Table S2 of the supplementary material. We used a Veriti™ thermal cycler (Applied Biosystems, Foster City, CA, USA) to perform the PCR reaction. Amplification products were visualized in an ABI-3500 Sanger sequencer at Escuela de Biología at Universidad de Costa Rica, using Hi-Di™ Formamide and GeneScan™ 500 LIZ™ dye Size Standard (Applied Biosystems). Genotypes were scored manually using GeneMarker Software version 2.6.4 (SoftGenetics).

### Clonal assignment

We used GenoDive v. 3.06<sup>53</sup> to determine the number of genets and associated ramets in each population. For *V. pompona*, the dataset included 146 individuals and 10 of 11 loci (mVplCIR016 was excluded due to missing data). For *V. odorata*, the dataset comprised 75 individuals and all 10 loci. A Stepwise Mutation Model was applied in both analyses, excluding missing data from the counts. For *V. pompona*, the minimum genetic distance





**Fig. 4.** A map of Costa Rica showing the locations of *V. odorata* (triangles) and *V. pompona* (circles) populations. The colors grouping the circles in *V. pompona* correspond to the five regions in which the populations are categorized: Blue = North Pacific; Pink = Central Pacific - Carara; Purple = Central Pacific - Puriscal, Yellow = South Pacific - Térraba Green = South Pacific - Peninsula. The numbers in brackets indicate the number of individuals collected at each site. In the zoomed-in view of the Piro Biological Station, the names in light orange represent *V. odorata* populations, while the names in light green represent *V. pompona* populations. Abbreviations are as follows: VT = Vanilla Trail, PST = Piro Station Trail, PPT = Pica Pica Trail, STH = Stairway to Heaven. The map was created using QGIS software (version 3.30.3-<sup>s</sup>-Hertogenbosch, available at <https://qgis.org/>), with minor modifications made in Adobe Photoshop<sup>®</sup> CC (Adobe Systems Inc., California, U.S.A.). QGIS is licensed under the GNU General Public License (<https://www.gnu.org/licenses>).

threshold for distinguishing different clonal lineages was established at 1, whereas for *V. odorata*, the threshold was set at 0<sup>54</sup> and the “Make clones specific to every population” option was selected for the analyses in both species. These thresholds were estimated from a frequency distribution of genetic distances, assuming that the first peak in the histogram was likely due to somatic mutations or genotyping errors<sup>55</sup>. The threshold was chosen between the first peak and the next increase in genetic distance frequencies. To assess clonal structure, we tested the probability of observing the detected clonal diversity by selecting the number of genotypes as the statistic for both species. We conducted 9,999 permutations, with randomization set to “alleles over individuals within populations” for *V. pompona* and “randomize alleles over individuals across all populations” for *V. odorata*.

We estimated the number of genets and calculated the clonality percentage by dividing the number of genets by the total number of individuals in each population. We also calculated Simpson’s Diversity Index (SDI), which measures genetic diversity by quantifying the probability that two randomly selected individuals belong to different genotypes. Additionally, we assessed evenness (E), which indicates the uniformity of genotype distribution among sub-populations. An evenness value of 1 indicates that all genotypes are equally represented, implying balanced diversity, whereas lower values indicate an uneven distribution, with some genotypes dominating. All these estimates were calculated for each population using GenoDive v. 3.06<sup>53</sup>. The analysis involved a bootstrap test for differences in clonal diversity with 9999 permutations and a jackknife procedure with increasing sample sizes, using 1000 permutations.

## Genetic diversity and structure

Genetic variation and its distribution within populations of *V. pompona* and *V. odorata* were described by estimating the average number of alleles (A), allelic richness to compensate for differences in sampling size per population (Ar), effective number of alleles (Ae), average observed ( $H_O$ ) and expected ( $H_e$ ) heterozygosity, and the inbreeding coefficient ( $F_{IS}$ ). We tested each locus for significant deviations from Hardy-Weinberg equilibrium (HWE) using the exact test implemented in the Genepop software<sup>56</sup>. We estimated the significance of inbreeding coefficients using 1000 bootstrap samples via the *mmod* package<sup>57</sup>. To evaluate differences in allele frequencies among populations, we calculated the Nei's  $G_{ST}$  index<sup>58</sup>. Isolation-by-distance (IBD) was measured by comparing Euclidean distances between populations using Nei's  $G_{ST}$  and the significance of IBD was determined using a Mantel test with 9,999 permutations. All genetic diversity and structure estimates were calculated using the *adegenet*<sup>59</sup> and *poppr*<sup>54,60</sup> packages in R version 4.2.3<sup>61</sup>. All genetic diversity estimates were calculated in two separate analyses, one including clones (IC) and a second analysis excluding clones (EC) within each population (see sub-section Clonal assignment).

We visualized the dissimilarity among individuals and populations using a Discriminant Analysis of Principal Components (DAPC) as implemented in the *adegenet* package<sup>59</sup>. Cross-validation was conducted by partitioning the data into a training set (75%) and a validation set (25%), repeating this process 1,000 times at each step of Principal Component (PC) retention. To determine the optimal number of PC axes, we used the *xvalDapc()* function from the *adegenet* package with 30 repetitions. We then refined our search to the five PC axes closest to the previously selected PC, increasing the number of replicates to 999. We retained 25 components and three discriminant functions for *V. pompona* DAPC, while in *V. odorata* we retained seven components and three discriminant functions.

## Fine scale spatial genetic structure (FSGS)

Autocorrelation analyses based on multilocus pairwise kinship coefficients<sup>47</sup> were calculated using SpaGeDi 1.5<sup>62</sup>. We selected distance classes with at least a minimum of 30 pairwise comparisons per distance class<sup>63</sup>. Fine-scale spatial genetic structure (FSGS) analyses were conducted globally for each species, combining all sampled populations into a single analysis per species. These analyses were conducted both with clones included (IC) and excluded (EC), which resulted in different distance intervals for each set of analyses. In *V. odorata*, the IC dataset had 19 distance classes, while the EC dataset had seven. For *V. pompona*, 33 distance classes were defined in the IC dataset, compared to 15 in the EC dataset. We also ran separate FSGS analyses on each population to see if patterns were consistent across populations.

Using the IC datasets, we calculated the coancestry coefficient  $\rho_{ij}$  to measure FSGS within populations that had at least 15 individuals<sup>47</sup>. For *V. pompona*, individual analyses were performed on six populations (VT, STH, La Pavona, Los Chocuacos, Clavera, and San Rafael), while for *V. odorata*, only three populations (PST, VT, and PPT) were analyzed. Permutation tests were conducted to evaluate the statistical significance of the observed patterns, with 10,000 permutations assigned to each test.

## Data availability

The SSR genotypes are available in the Zenodo repository at <https://zenodo.org/records/14195783>.

Received: 21 November 2024; Accepted: 26 May 2025

Published online: 02 June 2025

## References

- Givnish, T. et al. (ed, J.) Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc. Royal Soc. B: Biol. Sci.* **282** 20151553 (2015).
- Karremans, A. P. et al. A reappraisal of Neotropical *Vanilla*. With a note on taxonomic inflation and the importance of alpha taxonomy in biological studies. *Lankesteriana* **20**, 395–497 (2020).
- Besse, P. et al. RAPD genetic diversity in cultivated vanilla: *Vanilla planifolia*, and relationships with *V. tahitensis* and *V. pompona*. *Plant Sci.* **167**, 379–385 (2004).
- Loeillet, D. The international vanilla market: Price is the main handicap (2003).
- Chambers, A. et al. Genotyping-By-Sequencing diversity analysis of international *Vanilla* collections uncovers hidden diversity and enables plant improvement. *Plant Sci.* **311**, 111019 (2021).
- Bory, S. et al. Development of microsatellite markers in cultivated vanilla: polymorphism and transferability to other vanilla species. *Sci. Hortic.* **115**, 420–425 (2008).
- Minoo, D. et al. Genetic variations and interrelationships in *Vanilla planifolia* and few related species as expressed by RAPD polymorphism. *Genet. Resour. Crop Evol.* **55**, 459–470 (2008).
- Schlüter, P. M., Arenas, M. A. S. & Harris, S. A. Genetic variation in *Vanilla planifolia* (Orchidaceae). *Econ. Bot.* **61**, 328–336 (2007).
- Hu, Y. et al. Genomics-based diversity analysis of *Vanilla* species using a *Vanilla planifolia* draft genome and Genotyping-By-Sequencing. *Sci. Rep.* **9**, 1–16 (2019).
- Martinez, R. C. & Engleman, E. M. Caracterización de dos Tipos de *Vanilla planifolia*. *Acta Bot. Mex* **25**, 49–59 (1993).
- Watteyn, C. et al. Vanilla distribution modeling for conservation and sustainable cultivation in a joint land sparing/sharing concept. *Ecosphere* **11** (2020).
- Verma, P. C. et al. The extent of genetic diversity among *Vanilla* species: comparative results for RAPD and ISSR. *Ind. Crops Prod.* **29**, 581–589 (2009).
- Simiyu, L. N., Wolukau, J. N. & Oyoo, M. E. Diversity assessment of vanilla (*Vanilla* species) accessions in selected counties of Kenya using simple sequence repeats (SSRs) markers. *Afr. J. Biotechnol.* **19**, 736–746 (2020).
- Gigant, R. L. et al. Combining pollination ecology and fine-scale Spatial genetic structure analysis to unravel the reproductive strategy of an insular threatened Orchid. *South. Afr. J. Bot.* **105**, 25–35 (2016).
- Andriamihaja, C. F., Ramarosandratana, A. V., Grisoni, M., Jeannoda, V. H. & Besse, P. Drivers of population divergence and species differentiation in a recent group of Indigenous orchids (*Vanilla* spp.) in Madagascar. *Ecol. Evol.* **11**, 2681–2700 (2021).
- Gigant, R. L. et al. Microsatellite markers confirm self-pollination and autogamy in wild populations of *Vanilla mexicana* Mill.(syn. *V. inodora*)(Orchidaceae) in the island of Guadeloupe. *Microsatellite Markers*. **10**, 64674 (2016).

17. Gigant, R. et al. Nineteen polymorphic microsatellite markers from two African *Vanilla* species: across-species transferability and diversity in a wild population of *V. humblotii* from Mayotte. *Conserv. Genet. Resour.* **4**, 121–125 (2012).
18. Hamrick, J. L. & Godt, M. J. W. Effects of life history traits on genetic diversity in plant species. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **351**, 1291–1298 (1996).
19. Childers, N. F. & Cibes, H. R. *Vanilla* culture in Puerto Rico (1948).
20. Weiss, E. A. Chapter 7: Orchidaceae. *Spice Crops* 136–154 (2002).
21. Bory, S., Grisoni, M., Duval, M. F. & Besse, P. Biodiversity and preservation of vanilla: present state of knowledge. *Genet. Resour. Crop Evol.* **55**, 551–571 (2008).
22. Soto-Arenas, M. A. Filogeografía y recursos genéticos de las vainillas de México (1999).
23. Pemberton, R. W., Wheeler, G. S., Madeira, P. T. & Bee (Hymenoptera: Apidae) pollination of *Vanilla planifolia* in Florida and their potential in commercial production. *Fla. Entomol.* **106**, 230–237 (2023).
24. Quezada-Euán, J. J. G., Guerrero-Herrera, R. O., González-Ramírez, R. M. & MacFarlane, D. W. Frequency and behavior of *Melipona* stingless bees and Orchid bees (Hymenoptera: Apidae) in relation to floral characteristics of vanilla in the Yucatán region of Mexico. *PLoS One.* **19**, e0306808 (2024).
25. Favre, F. et al. A genome-wide assessment of the genetic diversity, evolution and relationships with allied species of the clonally propagated crop *Vanilla planifolia* jacks. Ex Andrews. *Genet. Resour. Crop Evol.* **69**, 2125–2139 (2022).
26. Epperson, B. K. & Clegg, M. T. Spatial-autocorrelation analysis of flower color polymorphisms within substructured populations of morning glory (*Ipomoea purpurea*). *Am. Nat.* **128**, 840–858 (1986).
27. Antonovics, J. & Levin, D. A. The ecological and genetic consequences of density-dependent regulation in plants. *Annu. Rev. Ecol. Syst.* **11**, 411–452 (1980).
28. Doust, L. L. Population dynamics and local specialization in a clonal perennial (*Ranunculus repens*): I. The dynamics of ramets in contrasting habitats. *J. Ecol.* **743**–755 (1981).
29. Yun, S. A., Son, H. D., Im, H. T. & Kim, S. C. Genetic diversity and population structure of the endangered Orchid *Pelatantheria scolopendrifolia* (Orchidaceae) in Korea. *PLoS One.* **15**, e0237546 (2020).
30. Charpentier, A. Consequences of clonal growth for plant mating. *Evol. Ecol.* **15**, 521–530 (2001).
31. Eckert, C. G. The loss of sex in clonal plants. In *Ecology and Evolutionary Biology of Clonal Plants: Proceedings of Clone-2000. An International Workshop Held in Obergurgl, Austria, 20–25 August 2000*. 279–298 (2002).
32. Judson, O. P. & Normark, B. B. Ancient asexual scandals. *Trends Ecol. Evol.* **11**, 41–46 (1996).
33. Balloux, F., Lehmann, L. & de Meeüs, T. The population genetics of clonal and partially clonal diploids. *Genetics* **164**, 1635–1644 (2003).
34. Brzosko, E., Wróblewska, A. & Ratkiewicz, M. Spatial genetic structure and clonal diversity of Island populations of Lady's slipper (*Cypripedium calceolus*) from the Biebrza National park (northeast Poland). *Mol. Ecol.* **11**, 2499–2509 (2002).
35. Andriamihaja, C. F., Ramarosandratana, A. V., Grisoni, M., Jeannoda, V. & Besse, P. The leafless *Vanilla* species-complex from the South-West Indian ocean region: A taxonomic puzzle and a model for Orchid evolution and conservation research. *Divers. (Basel)*. **12**, 443 (2020).
36. Chung, M. Y., Nason, J. D. & Chung, M. G. Implications of clonal structure for effective population size and genetic drift in a rare terrestrial Orchid, *Cremastra appendiculata*. *Conserv. Biol.* **18**, 1515–1524 (2004).
37. Orive, M. E. Effective population size in organisms with complex life-histories. *Theor. Popul. Biol.* **44**, 316–340 (1993).
38. Phillips, R. D., Dixon, K. W. & Peakall, R. Low population genetic differentiation in the orchidaceae: implications for the diversification of the family. *Mol. Ecol.* **21**, 5208–5220 (2012).
39. Ackerman, J. D. Coping with the epiphytic existence: Pollination strategies. *Selbyana* 52–60 (1986).
40. Karremans, A. P. et al. First evidence for multimodal animal seed dispersal in orchids. *Curr. Biol.* **33**, 364–371 (2023).
41. Kimura, M. & Weiss, G. H. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **49**, 561 (1964).
42. Fuchs, E. J. & Hamrick, J. L. Mating system and pollen flow between remnant populations of the endangered tropical tree, *Guaiaicum sanctum* (Zygophyllaceae). *Conserv. Genet.* **12**, 175–185 (2011).
43. Watteyn, C. et al. Trick or treat? Pollinator attraction in *Vanilla pompona* (Orchidaceae). *Biotropica* **54**, 268–274 (2022).
44. Adler, G. H. Fruit and seed exploitation by central American spiny rats, *Proechimys semispinosus*. *Stud. Neotrop. Fauna Environ.* **30**, 237–244 (1995).
45. Soto-Arenas, M. A. & Dressler, R. L. A revision of the Mexican and central American species of *Vanilla* plumier ex miller with a characterization of their ITS region of the nuclear ribosomal DNA. *Lankesteriana* **9**, 285–354 (2010).
46. Vekemans, X. & Hardy, O. J. New insights from fine-scale Spatial genetic structure analyses in plant populations. *Mol. Ecol.* **13**, 921–935 (2004).
47. Loiselle, B. A., Sork, V. L., Nason, J. & Graham, C. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am. J. Bot.* **82**, 1420–1425 (1995).
48. Karremans, A. P., Watteyn, C., Scaccabarozzi, D., Pérez-Escobar, O. A. & Bogarín, D. Evolution of seed dispersal modes in the orchidaceae: has the Vanilla mystery been solved? *Horticulturae* **9**, 1270 (2023).
49. Barreda-Castillo, J. M., Menchaca-García, R. A., Pérez-Silva, A., Sánchez-Coello, N. G. & Luna-Rodríguez, M. Influencia de La temperatura En La infectividad de F.sarium oxysporum F. Sp. vanillae En Vanilla planifolia y En Híbridos V. planifolia X V. pompona. *Biotechnia* **25**, 177–183 (2023).
50. Barreda-Castillo, J. M. et al. Morphological and physio-chemical responses to PEG-induced water stress in *Vanilla planifolia* and *V. pompona* hybrids. *Int. J. Mol. Sci.* **24**, 4690 (2023).
51. Menchaca, R. A. et al. Germinación in vitro de Híbridos de *Vanilla planifolia* y *V. pompona*. *Rev. Colomb Biotechnol.* **13**, 80–84 (2011).
52. Doyle, J. J. Isolation of plant DNA from fresh tissue. *Focus (Madison)*. **12**, 13–15 (1990).
53. Meirmans, P. G. Genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Mol. Ecol. Resour.* **20**, 1126–1131 (2020).
54. Kamvar, Z. N. & Grünwald, N. J. *Analysis of Multilocus Genotypes and Lineages in poppr packageVersion ('poppr')*.
55. Meirmans, P. G. & Van Tienderen, P. H. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes*. **4**, 792–794 (2004).
56. Raymond, M. & Rousset, F. *GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism* (1995).
57. Winter, D. J. MMOD: an R library for the calculation of population differentiation statistics. *Mol. Ecol. Resour.* **12**, 1158–1160 (2012).
58. Nei, M. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci.* **70**, 3321–3323 (1973).
59. Jombart, T. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).
60. Kamvar, Z. N., Tabima, J. F. & Grünwald, N. J. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**, e281 (2014).
61. Team, R. C. R: A language and environment for statistical computing; Preprint (2018).
62. Hardy, O. J. & Vekemans, X. SPAGeDi: a versatile computer program to analyse Spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes*. **2**, 618–620 (2002).
63. Waser, N. M. & Mitchell, R. J. Nectar standing crops in *Delphinium nelsonii* flowers: Spatial autocorrelation among plants? *Ecology* **71**, 116–123 (1990).



## Acknowledgements

This work was funded by Projects C0049 and C3464 of the Vicerrectoría de Investigación of the Universidad de Costa Rica, as well as the Sistema de Estudios de Posgrado and the Graduate Thesis Support Fund from the Vicerrectoría de Investigación. We thank Charlotte Watteyn, Marco Cedeño Fonseca, and Mauricio Fernández Otárola for their field assistance. We also acknowledge Ruth Madrigal-Brenes for her guidance and support in the laboratory. This manuscript was completed during E.J.F.'s sabbatical leave.

## Author contributions

M.A.S.-S., E.J.F., and A.P.K. designed the study and collected the samples. M.A.S.-S. conducted the lab work, as well as the statistical and genetic analyses. M.A.S.-S. drafted the manuscript. E.J.F. provided thorough revisions to the manuscript. All authors contributed to, completed, and approved the final manuscript.

## Declarations

### Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-04101-5>.

**Correspondence** and requests for materials should be addressed to M.A.S.-S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025